

Identificação automatizada de textos em imagens com Amazon Rekognition

Automated identification of text in images with Amazon Rekognition

Identificación automatizada de texto en imágenes con Amazon Rekognition

Recebido: 19/02/2023 | Revisado: 27/02/2023 | Aceitado: 03/03/2023 | Publicado: 14/03/2023

Jardel Silas da Silva Barata

ORCID: <https://orcid.org/0000-0003-4945-3847>
Universidade do Estado do Pará, Brasil
E-mail: jardelsilvati@gmail.com

Lucas Ravele De Sousa Teixeira

ORCID: <https://orcid.org/0000-0002-4571-2912>
Universidade do Estado do Pará, Brasil
E-mail: lucasravelly7@gmail.com

Bruno Campos Da Silva

ORCID: <https://orcid.org/0000-0003-2266-6981>
Universidade do Estado do Pará, Brasil
E-mail: brunocamposdasilva81@gmail.com

Thiago Nicolau Magalhães de Souza Conte

ORCID: <https://orcid.org/0000-0002-1288-366X>
Universidade do Estado do Pará, Brasil
E-mail: thiagonconte@uepa.br

Wilker José Caminha dos Santos

ORCID: <https://orcid.org/0000-0002-5265-583X>
Universidade do Estado do Pará, Brasil
E-mail: wilkercaminha@uepa.br

Resumo

O reconhecimento de texto em imagens é um desafio na área de visão computacional devido à variedade de fontes, qualidade da imagem, tamanho e cores presentes nas imagens. Nesse contexto, este trabalho tem como objetivo desenvolver uma aplicação de reconhecimento de textos em imagens utilizando a API Amazon Rekognition e avaliar sua precisão. Para isso, é proposto um algoritmo baseado em técnicas de deep learning capaz de alcançar uma precisão acima de 90% na localização e extração de texto nas imagens, utilizando métodos de extração de dados da função de detecção de texto da API Amazon Rekognition. Este artigo também tem o potencial de contribuir para o avanço de trabalhos futuros no campo da visão computacional, com ênfase na detecção de texto em imagens. Por fim, o estudo conclui que a API de detecção de texto da Amazon Rekognition é relevante na análise de dados, considerando que é treinada com grandes quantidades de dados de imagem para aprender características relevantes, alcançando uma precisão acima de 90%. No entanto, é necessário considerar que a qualidade da imagem e o tipo de fonte utilizada podem influenciar na precisão dos resultados.

Palavras-chave: Amazon Rekognition; Deep learning; Reconhecimento de texto; Análise de dados.

Abstract

The recognition of text in images is a challenge in the field of computer vision due to the variety of sources, image quality, size, and colors present in images. In this context, this work aims to develop an application for recognizing text in images using the Amazon Rekognition API and evaluate its accuracy. To achieve this, an algorithm based on deep learning techniques is proposed, capable of achieving an accuracy above 90% in the location and extraction of text in images, using data extraction methods from the text detection function of the Amazon Rekognition API. This article also has the potential to contribute to the advancement of future work in the field of computer vision, with a focus on text detection in images. Finally, the study concludes that the text detection API of Amazon Rekognition is relevant in data analysis, considering that it is trained with large amounts of image data to learn relevant characteristics, achieving an accuracy above 90%. However, it is necessary to consider that image quality and the type of font used can influence the accuracy of the results.

Keywords: Amazon Rekognition; Deep learning; Text recognition; Data analysis.

Resumen

El reconocimiento de texto en imágenes es un desafío en el campo de la visión computacional debido a la variedad de fuentes, calidad de imagen, tamaño y colores presentes en las imágenes. En este contexto, el objetivo de este trabajo es desarrollar una aplicación para reconocer texto en imágenes utilizando la API Amazon Rekognition y evaluar su precisión. Para lograrlo, se propone un algoritmo basado en técnicas de deep learning capaz de lograr una precisión superior al 90% en la localización y extracción de texto en las imágenes, utilizando métodos de extracción de datos de

la función de detección de texto de la API Amazon Rekognition. Este artículo también tiene el potencial de contribuir al avance de futuros trabajos en el campo de la visión computacional, con un enfoque en la detección de texto en imágenes. Finalmente, el estudio concluye que la API de detección de texto de Amazon Rekognition es relevante en el análisis de datos, considerando que está entrenada con grandes cantidades de datos de imagen para aprender características relevantes, alcanzando una precisión superior al 90%. Sin embargo, es necesario tener en cuenta que la calidad de la imagen y el tipo de fuente utilizada pueden influir en la precisión de los resultados.

Palabras clave: Amazon Rekognition; Aprendizaje profundo; Reconocimiento de texto; Análisis de datos.

1. Introdução

Reconhecimento de texto automático é um dos problemas mais difíceis em visão computacional. Embora muitos métodos de detecção de texto foram estudados no passado, o problema permanece em aberto (Romero, 2018). Segundo Silva (2019) os problemas de reconhecimento de textos acompanham pesquisadores da área de visão computacional há anos, e encontrava-se uma série de análises a respeito de técnicas voltadas à Optical Character Recognition (OCR), evoluções, dificuldades em criar uma solução que efetivamente faça um reconhecimento eficiente de textos.

Com o rápido desenvolvimento dos dispositivos digitais, as imagens tornaram-se meios de comunicação populares no nosso dia a dia. De acordo com um estudo realizado por Zhao et al., (2010) as imagens podem acarretar diversos tipos de informações escondidas, entre essas estão os textos, que muitas vezes incorporados em imagens e vídeos, contêm muitas informações semânticas úteis para a compreensão e recuperação de informações. no entanto a grande diversificação de fontes, tamanho e cores torna o processo de detecção um pouco complexo, já que os algoritmos processam os metadados.

O objetivo do trabalho é desenvolver uma aplicação que faça o reconhecimento de textos em imagens utilizando a Amazon Rekognition API na sua ferramenta de detecção de textos em imagens e avaliar sua acurácia. A acurácia no presente trabalho é descrita como: (proporção de palavras extraídas e reconhecidas corretamente em uma imagem, em relação ao total de palavras retiradas) o algoritmo utiliza uma imagem específica como entrada e gera uma saída probabilística como resultados (Rosa et al., 2019).

A Amazon Web Services (AWS) disponibiliza a ferramenta Rekognition para reconhecimento de imagens e vídeos. Ela permite que seja adicionado recursos facilmente às suas aplicações. Os primeiros conceitos que surgiram acerca do seu conteúdo, buscam ter uma visão ampla sobre o tema, focando especialmente na detecção de textos em imagens – que é uma das funcionalidades que o presente projeto em evidência oferece (Indla, 2021).

Amazon Rekognition em sua API de detecção de textos em imagens, pouco se encontra documentos, trabalhos sobre este tema. Assim Justificamos que o estudo sobre esta ferramenta se tornam importante para fomentar a pesquisa científica e promover o avanço de novas tecnologias nesse segmento. a tecnologia é a base que agrega novos elementos à ecologia de interação e comunicação ao permitir que às máquinas execute tarefas tradicionalmente desempenhadas pelos seres humanos. (Kaufman, 2020)

No entanto a API (Application Programming Interface) que faz parte dos serviços de inteligência artificial da Amazon, promove a interação do usuário com a plataforma de análises de dados a partir do envio de imagens ou vídeos, onde os dados são processados pelo algoritmo que utiliza aprendizagem profunda (deep learning), capaz de detectar objetos, pessoas, cenas e atividades, além de extrair textos e detectar qualquer conteúdo inapropriado. (AWS, 2020).

Na última década, houve um aumento expressivo no número de aplicações de Aprendizado de Máquina (AM) que utilizam técnica de Aprendizado Profundo (AP – Deep Learning). (Pina et al., 2019). Deep Learning é um subcampo do Aprendizado de Máquina que emprega modelos com múltiplos níveis de representação e abstração para compreender dados como imagens, áudio e texto. (Oliveira, 2020). Este trabalho propôs uma API que utiliza aprendizagem profunda para definir padrões de análise e processar imagens selecionadas, com o objetivo de obter resultados precisos em relação aos textos identificados nas imagens escolhidas. A acurácia do reconhecimento desses textos é um dos resultados propostos pela API. A

técnica na área da visão computacional que hoje melhor resolve esses desafios é o aprendizado profundo (deep learning) (Cozman, & Kaufman, 2022).

Este artigo está organizado da seguinte forma: seção II referencial teórico sobre o tema do artigo, seção III metodologia, a seção IV descreve os Resultados e Discussão, sessão V conclusão.

2. Referencial Teórico

Aguiar e Santos (2022) fazem um comparativo de desempenho entre metodologias de modelos de inteligência Artificial Supervisionada, destaca a importância da visão computacional na indústria 4.0. Essa importância é atribuída ao surgimento de novas tecnologias, como a digitalização, a internet das coisas, a realidade virtual e a realidade aumentada, que possibilitaram a criação de aplicações que combinam técnicas de visão computacional com inteligência artificial e aprendizado de máquina. E em como essas tecnologias pode contribuir na sociedade despertando interesse em diversos setores, como na indústria, em sistema de controle de qualidade, em órgãos de segurança pública, no aprimoramento de sistemas de navegação de veículos autônomos, no reconhecimento de objetos em aplicações de robótica e em aplicação de uso geral.

Opara (2019) realizou um estudo comparativo entre as quatro plataformas em nuvem: AWS, Microsoft Windows Azure, Google App Engine e IBM Cloud. Ele analisou diversas características como tipo de serviços, armazenamento, banco de dados, segurança, linguagens de programação, preços, virtualização, serviço móveis, internet das coisas, backup e recuperação de dados e interface do usuário. Os resultados mostraram que a AWS é indicada para grandes empresas devido à sua cobertura global, o Microsoft Windows Azure é ideal para startups e organizações que usam servidores Windows, o Google App Engine é a opção mais econômica para desenvolvedores de software e aplicativos baseado na web, e a IBM Cloud se destaca pelos serviços de virtualização e nuvem privada. Este estudo é útil para potenciais usuários, como pequenas e médias empresas, desenvolvedores iniciantes e grandes empresas na escolha da plataforma que atenda às suas necessidades.

O trabalho de Silva (2019) analisa a acurácia (o quociente entre a quantidade de palavras extraídas e reconhecidas corretamente de uma imagem e o total de imagens analisadas) da ferramenta Google Cloud Vision, utilizando imagens naturais na detecção de textos. International Conference on Document Analysis and Recognition (ICDAR). Uma importante conferência da comunidade de pesquisa em reconhecimento de textos que premia pesquisadores em análise de documentos e propõe desafios para inflar e motivar a pesquisa nessa área. No ano de 2017 a ICDAR teve como desafio a Competição de Leitura Robusta (Robust Reading Competition), Era organizada da seguinte forma: localização de palavras, localização e reconhecimento de palavras e reconhecimento de palavras (alvo deste trabalho).

A Pesquisa de Basso et al (2020) Apresenta um novo método não supervisionado para detecção de textos em imagens de sensoriamento remoto. Na pesquisa foram utilizadas 3 imagens de satélites para implementação do método que usa técnicas clássicas de processamento digital de imagens (PDI) como filtros de morfologia matemática e de suavização gaussiana e funções básicas disponíveis no Matlab para obtenção dos resultados.

A Tese de Romero (2018) Apresenta um sistema automático de reconhecimento de texto em imagens de IDs, usa diferentes algoritmos que auxilia na localização, extração e reconhecimento aplicados à extração de textos em documentos de identificação genéricos e por fim realiza uma prova de conceito do sistema em um ambiente real. O autor utiliza o algoritmo MSER como método para localização junto com melhorias no contraste e a informação nas bordas dos objetos para localizar os possíveis caracteres. O processo de extração textual passa por quatro etapas: Localização, Seleção, Extração e Reconhecimento para cumprir os devidos objetivos, localiza a imagem no texto, fazer umas seleções refinadas selecionando as regiões que realmente possuem caracteres, A etapa de extração é fundamental para o reconhecimento dos caracteres pelo sistema de Reconhecimento Óptico de Caracteres (OCR), e a última etapa busca responder “o que tá escrito na imagem?”

3. Metodologia

A Metodologia é uma disciplina que se dedica a estudar, compreender e avaliar os métodos disponíveis para pesquisa acadêmica. Ela examina, descreve e avalia técnicas de pesquisa para coleta e processamento de informações com o objetivo de resolver problemas e questões de investigação. (Prodanov & Freitas, 2013). O método utilizado é classificado como quantitativa, pois envolve a coleta e análise sistemática de dados com o objetivo de obter informações numéricas e quantitativas. Isso ocorre por meio do processamento de dados em imagens para a obtenção de informações e análise de resultados de maneira objetiva. Essa abordagem de pesquisa é baseada na medição, geralmente numérica, de um número limitado de variáveis objetivas, com ênfase na comparação de resultados e no uso extensivo de técnicas estatísticas. (Wainer et al., 2007).

O trabalho utiliza uma abordagem de pesquisa descritiva, na medida em que se descreve a natureza e o funcionamento da Amazon Rekognition API, sua capacidade de reconhecimento de texto e seus métodos de análise de imagens, além de avaliar a precisão da ferramenta. A pesquisa descritiva é caracterizada pelo levantamento de dados, e pode ser considerada como um passo prévio para identificar fenômenos que não são explicados pelas teorias existentes. (Wazlawick, 2010).

3.1 Desenvolvimento da aplicação com JavaScript no NodeJS

Foi desenvolvido uma aplicação que realiza chamadas ao SDK para JavaScript no NodeJS, a fim de utilizar os serviços da web através da API. Antes de utilizar o JavaScript no NodeJS, é necessário configurar o SDK, incluindo:

- A região da Amazon Web Services (AWS) onde os serviços serão solicitados;
- As credenciais que autorizam o acesso aos recursos do SDK.

Para manipular informações em formato JSON no projeto, foi utilizada a dependência "body-parser". Foi criado um objeto com "aws rekognition" e realizada uma chamada à operação, passando parâmetros em formato JSON, que são tratados pela API. Após a configuração das dependências, o SDK já pode realizar as chamadas aos serviços da AWS por meio de programação. Por fim, é necessário configurar as "ACCESS KEY", ou seja, as chaves de acesso para fazer a chamada de serviço no ambiente.

3.2 Protocolo da API

A aplicação foi desenvolvida utilizando as tecnologias JavaScript, Node.js e Express.js. Para iniciar, o modo estrito é definido e os módulos express, body-parser, http e aws-sdk são importados. Uma aplicação Express é criada e o body-parser é utilizado para analisar dados no formato JSON, após a extração de dados os mesmos são organizados e processados utilizando JSON para armazenamento da informação em formato texto (Antonio, 2019). A configuração do AWS foi carregada a partir de um arquivo 'config.json' e o módulo Multer é importado para auxiliar com o upload de arquivos. Um servidor HTTP foi criado com a aplicação Express e as rotas GET e POST são configuradas. A rota GET retorna uma mensagem para o cliente, enquanto a rota POST armazena dados JSON e um arquivo do cliente em um bucket S3 da Amazon. Por fim, o script escuta a porta 3000 e inicia o servidor, conforme a Figura 1.

Figura 1- Rota de conexão.

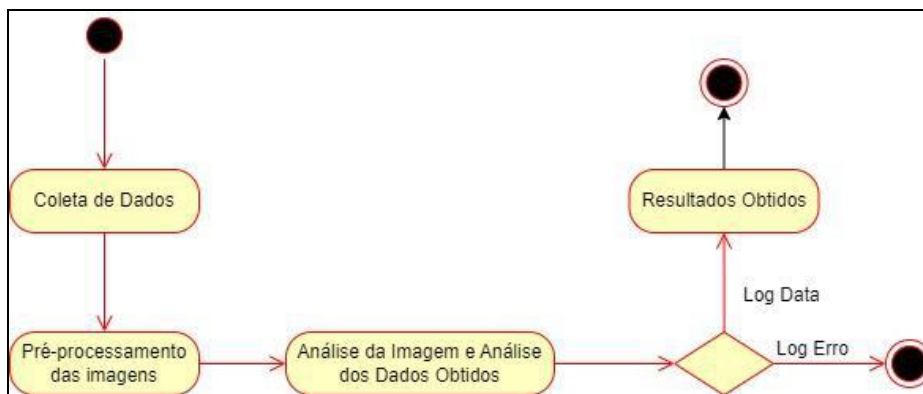


Fonte: Elaborada pelos Autores (2023).

A Figura 1 executa a rota de conexão do servidor local, para lidar com as solicitações de entrada do cliente.

A Figura 2 mostra uma ampla visão a seguir do algoritmo sendo o ponto de partida a coleta de dados, que se resume em enviar a imagem para a submissão das etapas seguintes, a imagem pode ser enviada em dois formatos, reconhecimento de texto e reconhecimento em Json. Neste passo a aplicação precisará carregar uma imagem para ser processada pela Amazon Regoknition, isso pode ser feito de várias maneiras, como carregar uma imagem a partir do computador ou bucket S3 da AWS.

Figura 2 – Diagrama de atividades.



Fonte: Elaborada pelos Autores (2023).

Na Figura 2 descreve o diagrama que representa as etapas que devem ser seguidas no processo do funcionamento da aplicação, vale lembrar que a API não armazena a imagem no banco de dado assim que é submetido outra imagem a anterior é apagada automaticamente. Esses dados são enviados através da API para ser analisada. O pré-processamento diz respeito a organização dos dados descarregados descritos na seção anterior com essa imagem é feito a coleta dos metadados utilizando AWS REKOGNITIONS para detecção de texto na imagem.

3.3 Pré-processamento

Após a coleta de dados e o pré-processamento das imagens a aplicação verifica se os dados JSON e os arquivos estão sendo recebidos corretamente pela rota POST, em seguida entra em uma condição. Se não haver erro na análise ele retorna os dados para o algoritmo, se houver erro é retornado uma mensagem no console, como é demonstrada na Figura 3

Figura 3 - Pré-processamento dos dados.

```
68     rekognition.detectText(params, function (err, data) {
69         if (err) {
70             console.log(err, err.stack); // um erro ocorreu
71         }
72         else {
73             console.log(data);           // resposta bem sucedida
```

Fonte: Elaborada pelos Autores (2023).

A Figura 3 apresenta um trecho do código que entra em um laço de condição, note que se houver erro a linha 70 retorna uma mensagem de erro, se não houver erro o algoritmo prossegue na leitura do código, como mostra a linha 73.

4. Resultados e Discussão

Ao processar uma imagem com o Amazon Rekognition, o algoritmo retorna informações detalhadas sobre o texto reconhecido na imagem. As informações são apresentadas em uma tabela, que inclui a palavra ou linha de texto reconhecida, o tipo de texto detectado ("LINE", "WORD", "CONFIDENCE") como é demonstrado na Figura 4. Confidence é a pontuação de confiança, é uma medida numérica que indica o nível de confiança do Amazon Rekognition na precisão da detecção de texto e dos pontos de geometria ao redor do texto. Quanto maior a pontuação de confiança, maior é a confiança do Amazon Rekognition na precisão dessas detecções.

Figura 4 -Aferição de nível de confiança .

```
77     table += "<tr>";
78     table += "<td>"+data.TextDetections[i].DetectedText+"</td>"
79     table += "<td>"+data.TextDetections[i].Type+"</td>"
80     table += "<td>"+data.TextDetections[i].Confidence+"</td>"
81     table += "</tr>"
```

Fonte: Elaborada pelos Autores (2023).

A Figura 4 mostra o trecho do código que gera a tabela com seus respectivos resultados, como é demonstrado na Figura 7.

O Amazon Rekognition foi projetado para reconhecer vários idiomas, incluindo inglês, francês, espanhol, alemão, italiano, português, russo, chinês, japonês, entre outros. A precisão da identificação de idioma pode variar de acordo com a qualidade da imagem e a complexidade da fonte. Alguns idiomas menos comuns ou antigos podem não ser reconhecidos com precisão pelo Amazon Rekognition.

Amazon Rekognition está disponível através de uma API RESTful, o que significa que pode ser facilmente integrada às suas aplicações, e é compatível com diversas linguagens de programação, tais como Java, Python, Ruby, entre outras.

4.1 Análise de Imagens

Note que na linha 4 na Figura 5 é criada a variável arquivo que vai receber a imagem em formato JPEG ou PNG, em seguida é criada outra variável na linha 11 chamada params, esta variável da linha 11 armazena a imagem e a transforma em formato json para ser analisada pela chamada da aws Rekognition como descrito na linha 16 da Figura 5.

Figura 5 – Estrutura da solicitação da detecção de texto da Amazon Rekognition.

```
1 //Codigo para trazer todas as informações do ARRAY em formato de TABELA (DetectedText - Type - Confidence)
2 app.post("/reconhecerTexto", upload.single('myfile_texto'), (request, response) => {
3 //arquivo que esta vindo do upload (MULTER)
4 var arquivo = request.file.buffer;
5 /* tem que ser a primeira linha de qualquer chamada do sdw para ele entender que tem que ser usada definindo regioao
6 em (formato json) regioao da virginia */
7 aws.config.update({region: 'us-east-1'});
8 // criando um objeto Rekognition
9 var rekognition = new aws.Rekognition();
10 // Paramentos - objeto parametro sempre em formato json
11 var params = {
12   Image: {
13     Bytes: arquivo
14   }
15 };
16 rekognition.detectText(params, function (err, data) {
17   if (err) {
18     console.log(err, err.stack); // um erro ocorreu
19   }
20   else {
21     console.log(data); // resposta bem sucedida
22     var table = "<table border=1>" // tabela html
23     for (var i=0; i < data.TextDetections.length; i++){
24       table += "<tr>";
25       table += "<td>"+data.TextDetections[i].DetectedText+"</td>"
26       table += "<td>"+data.TextDetections[i].Type+"</td>"
27       table += "<td>"+data.TextDetections[i].Confidence+"</td>"
28       table += "</tr>"
29     }
30     table += "</table>"
31     response.send(table);
32   }
33 });
34 });
35 });
```

Fonte: Elaborada pelos Autores (2023).

Após o processamento feito pela Amazon Rekognition, é dada como resposta, uma série de informações a respeito da imagem enviada, uma das informações mais relevantes para esse trabalho é a transcrição do texto e a localização do conteúdo textual na imagem como mostra a Figura 6.

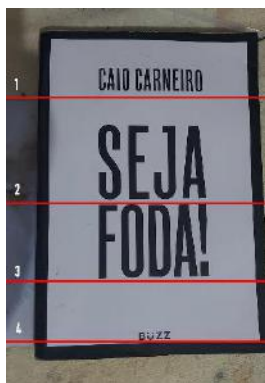
Figura 6– Estrutura da resposta de detecção de texto da Amazon Rekognition.

```
1 {
2   TextDetections: [
3     {
4       DetectedText: 'CAIO CARNEIRO',
5       Type: 'LINE',
6       Id: 0,
7       Confidence: 100,
8       Geometry: [Object]
9     },
10    {
11      DetectedText: 'SEJA',
12      Type: 'LINE',
13      Id: 1,
14      Confidence: 97.83336639404297,
15      Geometry: [Object]
16    },
17    {
18      DetectedText: 'FODAI',
19      Type: 'LINE',
20      Id: 2,
21      Confidence: 97.60860443115234,
22      Geometry: [Object]
23    },
24    {
25      DetectedText: 'BUZZ',
26      Type: 'LINE',
27      Id: 3,
28      Confidence: 80.35851287841797,
29      Geometry: [Object]
30    },
31    {
32      DetectedText: 'CAIO',
33      Type: 'WORD',
34      Id: 4,
35      ParentId: 0,
36      Confidence: 100,
37      Geometry: [Object]
38    },
39    {
40      DetectedText: 'CARNEIRO',
41      Type: 'WORD',
42      Id: 5,
43      ParentId: 0,
44      Confidence: 100,
45      Geometry: [Object]
46    },
47    {
48      DetectedText: 'SEJA',
49      Type: 'WORD',
50      Id: 6,
51      ParentId: 1,
52      Confidence: 97.83336639404297,
53      Geometry: [Object]
54    },
55    {
56      DetectedText: 'FODAI',
57      Type: 'WORD',
58      Id: 7,
59      ParentId: 2,
60      Confidence: 97.60860443115234,
61      Geometry: [Object]
62    }
63   ],
64   TextModelVersion: '3.0'
65 }
```

Fonte: Elaborada pelos Autores (2023).

Após o processamento da imagem o algoritmo retorna os resultados da análise da imagem, com a acurácia e nível de precisão através da exibição de uma tabela com os possíveis dados identificados e sua respectiva acurácia de confiança, como é observado na Figura 7.

Figura 7 – Estrutura da imagem de exemplo e resultado da detecção de texto da Amazon Rekognition.



CAIO CARNEIRO	LINE	100
SEJA	LINE	97.83336639404297
FODA!	LINE	97.60860443115234
BUZZ	LINE	80.35851287841797
CAIO	WORD	100
CARNEIRO	WORD	100
SEJA	WORD	97.83336639404297
FODA!	WORD	97.60860443115234
BUZZ	WORD	80.35851287841797

Fonte: Elaborada pelos Autores (2023).

Neste exemplo mostra como o algoritmo faz a leitura da imagem de linha por linha. O tamanho da fonte, o tipo, e a qualidade da imagem interfere na inspeção da imagem. Note que na linha 4 o nível de acurácia obtido no processamento da imagem ficou abaixo dos 90% devido o tipo da fonte.

5. Conclusão

Com os testes realizados concluímos que a Amazon Rekognition em sua API de detecção de texto torna-se bastante relevante na análise de dados pois a mesma é treinada com grandes quantidades de dados de imagem para aprender as características relevantes. Assim o algoritmo conseguiu atingir acima dos 90% da acurácia, com tudo alguns fatores como a qualidade da imagem e o tipo de fonte, pode dificultar na análise da imagem comprometendo a precisão dos resultados. Foi observado que o algoritmo encontrou dificuldades em analisar imagens de baixa qualidade e fontes especiais. Além disso, o processamento rápido da análise da imagem pode afetar a precisão da leitura. Os estudos ainda são novos nesse campo utilizando a tecnologia AWS, poucos é encontrado materiais de estudos e pesquisas científicas, para este tipo de pesquisa. Investigar sobre essas ferramentas é extremamente importante pois as mesmas tem o potencial de transformar muitos aspectos da nossa sociedade, podendo melhorar a eficiência e ajudar a automatizar tarefas repetitivas e complexas, tornando-as mais eficientes e precisas, e também contribuir para Desenvolvimento de novas tecnologias, como robótica, automação e aprendizado de máquina.

O presente trabalho deixa para atividades futuras um estudo mais detalhado sobre estas ferramentas: Microsoft Cognitive Services, ABBYY FlexiCapture e Google Cloud Vision. Ferramentas que são promissoras para o seu tempo. Dessa forma, espera-se que os trabalhos futuros possam contribuir para a expansão do conhecimento e uso dessas ferramentas, proporcionando resultados ainda mais satisfatórios para empresas e usuários em geral.

Referências

- Antonio, D. V. (2021). Implementação de protótipo baseado na tecnologia OCR aplicada ao reconhecimento de rótulos para busca em banco de dados. (Dissertação de mestrado, Universidade Federal do Rio de Janeiro).
- Aguiar, B., & Santos, G. (2022). Comparativo de desempenho entre metodologias de modelos de inteligência artificial supervisionada. [Comparative performance between supervised artificial intelligence model methodologies]. *Revista Brasileira de Tecnologia Aplicada*, 10(2), 45-58.
- AWS. (2020). Developer guide: Detecting text. Amazon Web Services. Retrieved from <https://docs.aws.amazon.com/rekognition/latest/dg/text-detection.html>.
- AWS. (2020). Amazon Rekognition. Retrieved from <https://aws.amazon.com/pt/rekognition/>
- Basso, D. P., Colnago, M., & Casaca, W. (2020). Um método não-supervisionado de detecção de oclusões textuais para imagens de sensoriamento remoto. *Proceeding Series of the Brazilian Society of Computational and Applied Mathematics*, 7(1).
- Cozman, F. G., & Kaufman, D. (2022). Viés no aprendizado de máquina em sistemas de inteligência artificial: a diversidade de origens e os caminhos de mitigação. *Revista USP*, 1(135), 195-210.
- Indla, R. K. (2021). An overview on Amazon Rekognition technology.
- Inteligência Artificial. (2013). (3ª ed.). Rua Sete de Setembro, 111 – 16 o andar: 1350.
- Kaufman, D. (2020). Inteligencia artificial: repensando a mediação. *Brazilian Journal of Development*, 6(9), 67621-67639.
- Oliveira, W. (Data não especificada). Identificar autoridades por meio de reconhecimento facial. Uso de tecnologia de visão computacional como alternativa para antigo processo de fotogramas (Carômetro).
- Opara, C. M. (2019). Cloud computing in Amazon Web Services, Microsoft Windows Azure, Google App Engine and IBM cloud platforms: A comparative study. (Doctoral dissertation). Near East University.
- Pina, D. B., Cruz, S., Ferreira, R. C., Silva, M. L., & Matos, D. M. (2019). Análise de hiperparâmetros em aplicações de aprendizado profundo por meio de dados de proveniência. In *Anais do XXXIV Simpósio Brasileiro de Banco de Dados* (pp. 223-228). SBC.
- Prodanov, C. C., & Freitas, E. C. de. (2013). Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico (2ª ed.). Novo Hamburgo: Universidade FEEVALE.
- Romero, R. V. (Data não especificada). Processo automático de reconhecimento de texto em imagens de documentos de identificação genéricos. Tese de Doutorado. Universidade de São Paulo.
- Rosa, A. L., Martins, V. S., Pires, A. M. A., Oliveira, A. R. A., & Araujo, R. B. (2019). Classificação de imagens de frutas utilizando aprendizado de máquina. *Anais do Congresso Brasileiro de Informática na Educação*, 30(1), 1-10.
- Silva, R. O. (2019). Análise de desempenho da Google Cloud Vision API em leitura de textos provenientes de imagens naturais.
- Santos, M. (2010). Protótipo para Mineração de Opinião em Redes Sociais: Estudo de Casos Seleccionados Usando o Twitter. Monografia. Departamento de Ciência da Computação, Universidade Federal de Lavras.
- Wainer, J., Borenstein, J., & Cukierman, H. L. (2007). Métodos de pesquisa quantitativa e qualitativa para a Ciência da Computação. Atualização em Informática, 1(221-262), 32-33.
- Wazlawick, R. S. (2010). Uma reflexão sobre a pesquisa em ciência da computação à luz da classificação das ciências e do método científico. *Revista de Sistemas de Informação da FSMA*, 6, 3-10.
- Zhao, M., Li, S., & Kwok, J. (2010). Text detection in images using sparse representation with discriminative dictionaries. *Image and Vision Computing*, 28(12), 1590-1599. doi: 10.1016/j.imavis.2010.03.015.